HPC AI500: A Benchmark Suite for HPC AI Systems

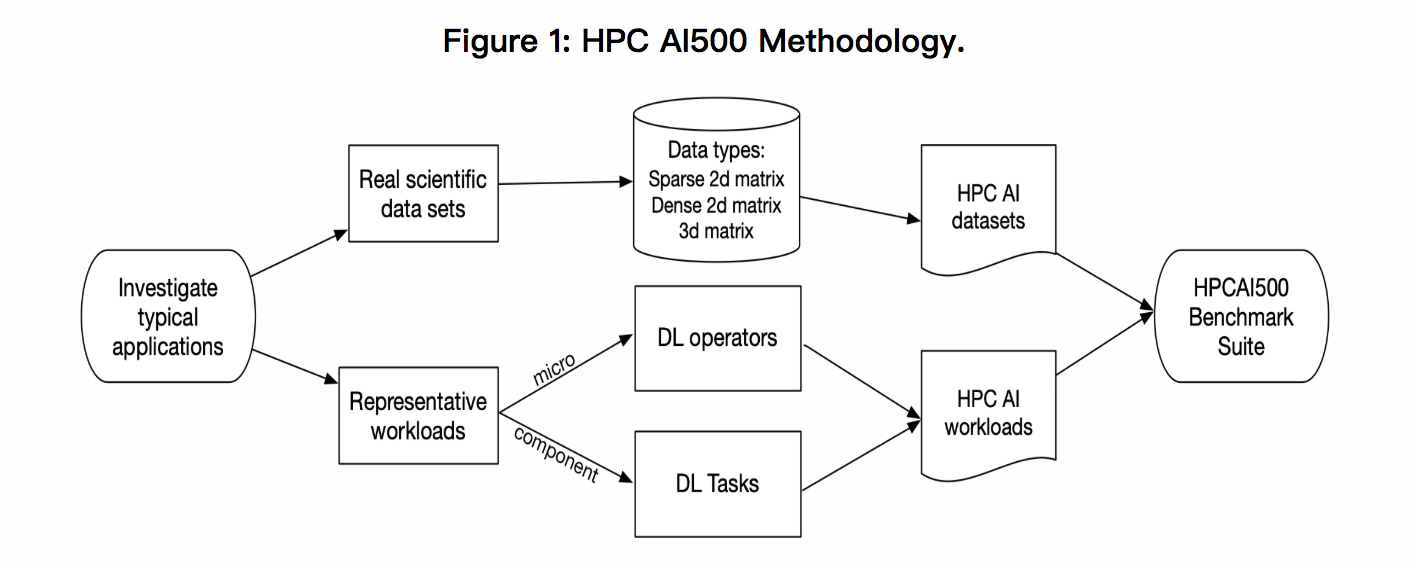
摘要：HPC上运行的科学DL应用越来越多，与传统高性能计算的IO特征有很多不同，因此需要新的benchmark去评估HPC系统跑科学DL应用的性能。本文就是提出了这样的一套benchmark。

1. 动机

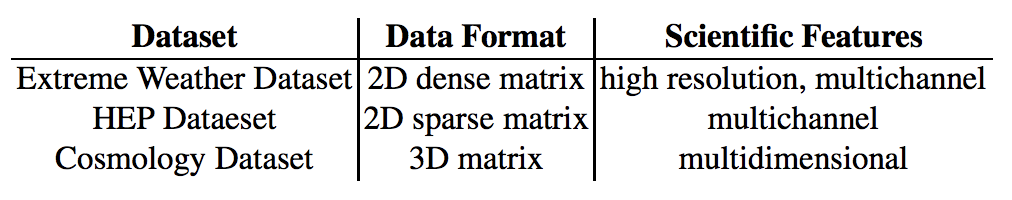
HPC上运行的科学DL应用越来越多，与传统高性能计算的IO特征有很多不同：大量的数据用高维矩阵表示、使用tensorflow等AI框架导致软件和硬件的协同设计困难、异构的计算平台非常复杂：CPU/GPU/Cambricon Diannan/Google TPU；因此次，需要新的benchamrk去衡量HPC系统跑科学DL的性能好坏；

1. 设计

具体的构造过程如下图：



这套benchmark包含两部分，其一是数据集，汇集了典型的不同的数据集，如下：



其二是能代表科学DL的AI workload；其中包含经典的神经网络（如ResNet、Faster-RCNN、DCGAN等），也包含很多基本的操作（如卷积、池化、全连接等，可以用来评估HPC系统性能的上限）

这套HPCAI500 benchmark有以下特点：

1. 代表性强：覆盖了代表性的科学领域，即气候分析，宇宙学，高能物理学，引力波物理学和计算生物学。
2. 评价指标考虑到准确度、性能、功耗、成本
3. 考虑应用程序场景，数据集和软件堆栈，选择18个科学DL benchmarks 具体参见http://www.benchcouncil.org/HPCAI500/index.html
4. 应用具有拓展性，可以在很多结点同时运行。